Antenna Parameter Measurement Network with Dual Attention and Focus Loss Using UAV

Ying Xu, Qirui Ke, Ziyi Jiang, Yikui Zhai, *Member, IEEE*, Angelo Genovese, *Member, IEEE*, Vincenzo Piuri, *Fellow, IEEE*, and Fabio Scotti, *Senior Member, IEEE*

Abstract—**Although deep learning has proven to be a successful and widely used technology across various industries, its drawbacks such as large models and difficulties in layout and maintenance in practical tasks have gradually become prominent. In view of the limitations and issues with the traditional method of measuring antenna parameters for Mobile Communication Base Stations (MCBS-APM), we are exploring the development of a new system that is designed to overcome the inefficiencies and potential risks associated with conventional labor-intensive methods. An effective measurement system which is composed of a Novel Instance Segmentation Network with Dual Attention and Focal Loss can accurately fathom out the antenna parameters in mobile communication base stations and completely subvert traditional measurement methods. To begin with, antenna video data is collected by unmanned aerial vehicle (UAV) which flies around the base station; then a designed instance segmentation network is employed to process and segment mobile communication base station antennas. At last, we implement real-time adjustments to control the actions of the UAV based on algorithmic measurements displayed on the accompanying mobile application. Our measurement system has been shown to greatly enhance measurement efficiency and accuracy, as evidenced by the results of our experiments. Quantitative results that are in line with industry standards show that our measurement system has strong robustness and reproducibility.**

Impact Statement — **The angle of down-tilt for mobile communication base station antennas is critical in ensuring optimal network coverage and communication quality. Traditionally, measuring the antenna declination parameters has been inefficient and poses danger to engineers as they have to climb base stations. We propose a fully automated MCBS-APM system using UAV. Our effective measurement system, which is composed of a Novel Instance Segmentation Network with Dual Attention and Focal Loss, represents a significant improvement over traditional methods MCBS-APM. With the results from the conducted experiments, it is fully proved that the proposed measurement system excels in improving efficiency and accuracy during measurement. Quantitative results in line with industry standards show that our measurement system has strong robustness and reproducibility.**

Index Tems*—*Antenna Parameters; Deep learning; Instance segmentation; Measurement system; UAV.

This work was supported by Key Research Projects for Universities of Guangdong Provincial Education Department (No. 2020ZDZX3031); Guang dong Basic and Applied Basic Research Foundation (No. 2021A1515011576); Fully Autonomous and Unattended AI Inspection Key Technology and Its Application in Public Security (No. 2021030103230006670); Guangdong, Hong Kong, Macao and the Greater Bay Area International Science and Technology Innovation Cooperation Project (No. 2021A050530080); Research and Application of Self-supervised Fusion Learning Method for Multi-modal Emotion Perception (No. 2022ZDZX1032); The key technology of fully autonomous UAV inspection and its application in intelligent land inspection (2220002000246); Development of Container Wisdom Panorama Trademark Quality Inspection System Based on Machine Vision (HX22105).

Ying Xu, Qirui Ke, Ziyi Jiang, Yikui Zhai are with the Department of Intelligent Manufacturing, Wuyi University, Jiangmen, China, 529020.

Angelo Genovese, Vincenzo Piuri, and Fabio Scotti are with the Departimento di Information, Universita, Degli Studi di Milano, via Celoria 18, 20133 Milano (MI), Italy.

Corresponding Author: Yikui Zhai. Email: yikuizhai@163.com.

I. INTRODUCTION

WITH the top spot of the competition occupied by deep learning technology again and again after the appearance of deep learning [1] by Hinton, the industry community has set off application researches on deep convolutional neural networks [2]. Under the premise that academia and industry promote mutual development, various new technologies have been proposed and verified with countless examples of successful applications in deep learning technology. More and more industries [3] have encountered bottlenecks in traditional methods for its difficulties to satisfy market demand. They tried to use these technologies which can make it easier and more efficient to change or even subvert traditional methods.

Optimal down-tilt angle is a vital consideration that impacts network coverage and communication quality. The down-tilt angles are essential parameters for characterizing the orientation of an antenna in mobile communication networks. Adjusting these angles is crucial in the network optimization process of mobile communication systems. The down-tilt angle is the angle between the antenna and the vertical plane. This industry is suffering from problems such as insufficient efficiency and uncontrollable risks of traditional methods. In response to these problems, some researchers proposed methods relying on expensive hardware equipment. On the basis of kinematic structure, an evaluating system without contact was developed by Xu et al [4] which can estimate the accurate posture through a small amount of antenna images as well as some simple operation conducted on the smartphone and though lacks accuracy with an error of 2°, which fails to satisfy the communication industry standard. In addition, Pedras et al. [5] proposed a novel quality of experience model that utilizes radio frequency channel indicators to estimate user-perceived quality. They determined the optimal antenna down-tilt angle by considering the average opinion score level and manually adjusting it accordingly. Undoubtedly, the

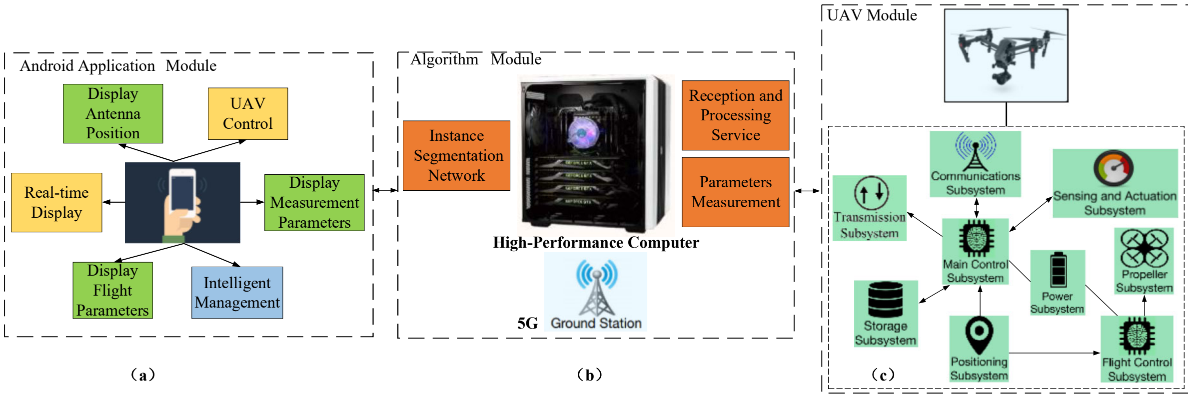


Fig. 1. The structure of our system. In particular, (a) highlights the key functional modules of the mobile application, such as UAV control and management, visualization of UAV flight parameters and antenna positions, and display of measurement results generated by the algorithm module. (b) presents the mechanism by which the UAV transmits its video stream to the high-performance algorithm server via the 5G network and generates real-time measurement results. Lastly, (c) delineates the functional modules of the UAV control subsystem.

contributions of the above methods should be acknowledged, and yet simultaneously their problems such as high labor costs, considerable measurement errors and demanding measurement environments are also worth noticing. Similarly, we develop a novel instance segmentation network with dual attention and focal loss for the system to measure the antenna down-tilt parameters for responding to these thorny problems. Firstly, UAV is utilized to fly around the base station to obtain antenna video data; then, 5G technology is adopted to transmit captured video data, and a designed instance segmentation network is employed to process captured video data and segment mobile communication base station antennas. Finally, Android Application (APP) displays algorithmic measurement results in real-time for the convenience of controlling UAV actions. Figure 1 can be referred to the architecture of the proposed system. The quantitative results that are in line with industry standards prove the strong robustness and reproducibility of our measurement system.

We conducted multiple empirical studies to showcase the efficacy of our system and compared it with existing methods to demonstrate its superiority as the most advanced approach. Our work contributes in the following ways:

1) This work proposes an efficient measuring system of antenna parameters to take the place of traditional methods for MCBS-APM that conduct with labors. Quantitative results that are in line with industry standards show that our measurement system has strong robustness and reproducibility.
2) A novel general-purpose instance segmentation network with a dual attention module placing channel attention before spatial attention is proposed. Self-adaptive refining the features is presented to produce prime mask of each antenna and take the measurement of the antenna parameter. And focal loss is adopted to reconcile the serious imbalance between negative and positive samples.
3) Our results indicate that the proposed method achieves superior performance in terms of speed and precision compared to existing state-of-the-art approaches..

The structure of the essay is as follows: Section II analyses the related works in how to address the relative problems, while Section III introduces the efficient measuring system of antenna parameters. Section IV provides an explanation of the experiment outcomes, and lastly Section V summarizes the work and expectations in this area.

II. Related Work

A. Instance Segmentation

Instance segmentation [6] occupying an essential, complicated and demanding position in computer vision, is able to localize different categories of object instances among different images for facilitating the prediction of the object instance-mask at specific pixel as well as the object class-label. Its application mostly lies in autonomous driving, surveillance and parameters measuring. With the appearance of Convolutional Neural Networks (CNN), scholars have proposed numerous instance segmentation frameworks, such as [7-10], where the rapid growth of the segmentation accuracy can be clearly observed. In [10], a series of experiments verified Mask R-CNN to be a direct as well as rapid instance segmentation approach. Fast/Faster R-CNN [11-12] has been shown to achieve superior results in predicting segmentation masks, compared to Fully Convolutional Network (FCN) [13], attributed to its use of box-regression and object classification. The Feature Pyramid Network (FPN) [14] can effectively capture stage-wise network features to enhance model performance. Specifically, the FPN utilizes a vertical network path and lateral connections to extract semantically rich features. The FPN architecture has shown impressive performance in numerous real-world scenarios, including autonomous driving, industrial inspection, and medical imaging, among others.

B. Attention Mechanism

Derived from deep learning, attention mechanism shares similarities with the attention mechanism of human vision in selecting the key information, ignoring other unimportant information and focusing on the important points among much information. With the proposal of [15], the effectiveness of the attention mechanism has been further proved, and it also opened a new era of attention mechanism research. Wang et al [16] introduced non-local operations that is the opening work of attention mechanism in computer vision, using self-attention mechanism to establish remote dependence. A module of dual attention mechanism, connecting the channel attention module as well as the spatial attention module by cascading was introduced by Woo et al [17]. This module enjoys wide applicability and can be integrated with any feed-forward convolutional neural network. Since context fusion fails to use the relationship between objects in the global view, Fu et al [18] raised a dual attention network based on a self-attention mechanism for seizing feature dependence in the spatial and channel dimension in respective manner. Specifically, double attention modules are attached to the dilated FCN so as to establish the sematic dependence model on the spatial and channel dimension respectively. As more and more research is conducted on attention mechanisms, their operation methods are becoming more and more diverse. But they all have a common core idea: to apply greater attention weight to the more interesting aspects of the feature.

C. Focal Loss

Focal Loss [19] specializes in analyzing a group of sparse hard examples, which can effectively protect against massive easy negatives flooding the detector in training. Romdhane et al [20] optimized a deep CNN model with a new loss function named focal loss, for favorable performance in detecting certain heartbeat categories (particularly for unbalanced data sets). The evaluation results also tell that assisted with focal loss function the classification accuracy of minority ethnic classification and the overall index can be improved. Building upon the application of focal loss in image-based object detection, [21] introduces an extension of binary cross entropy for point cloud-based object detection. Their evaluation demonstrates that focal loss yields the most significant improvement in 3 dimensional object detection compared to other examined hyperparameters. Li et al [22] proposed focal loss and channel attention in network design for the purpose of enhancing feature representation learning and improving tracking performance with attention to the interdependence of channels and the data imbalance between foreground and background. Focal loss can be adopted in multiple practical applications, mainly to improve unbalanced number of difficult and easy samples, and it can improve network performance in different computer vision tasks.

D. UAV

Results have proved that UAVs enjoy autonomy and flexibility in varied situations and tasks, so they can be applied in various fields. Samaras et al [23] introduced the application of counter UAV related tasks to many different sensors data and the improvement of deep learning after fusioning information with multi-sensors to facilatate the future application of counter UAV. Barbedo et al [24] turned to drones and deep learning technology to monitor the herd. The experiment confirms the feasibility of animal detection and determines the ideal ground sample distance. Furthermore, in [25] the trade-offs of deep convolutional neural network-based object detectors were discussed, highlighting their ability to enable UAVs to perform vehicle detection tasks in resource-constrained environments. According to Markov decision process and the partially observable one, Walker et al [26] designed a deep reinforcement learning framework for UAV indoor environment navigation. From all these fruitful achievements, it can be seen that UAVs with on-board cameras and computer vision functions of embedded systems widely enjoy increasing popularity in applications.

Algorithm 1:
Flowchart of the proposed system for MCBS-APM

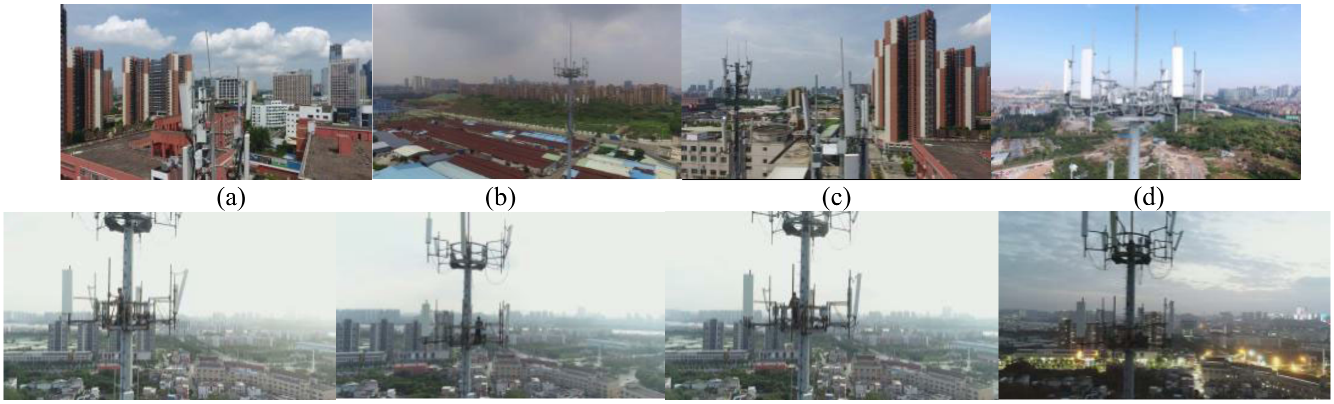
Input:
The latitude and longitude of the MCBS to be tested.
Output:
The APP interface displays the down-tilt angle d , area a , number of antennas n , and aspect ratio r of each antenna.
Step1: Enter the longitude and latitude of the MCBS to be tested in the APP;
Step2: UAV flies to the base station to be tested and orbits the base station according to preset parameters;
Step3: UAV captures antenna video and transmits it to the server in real time via high-bandwidth 5G network;
Step4: Algorithms in server process videos shot by UAV;
Step5: The number of antennas n is determined by the algorithms based on the pixel coordinates and threshold (with an empirical value of 50 for multiple verification);
Step6: Algorithms segment antennas and measure parameters (aspect ratio r , are a and down-tilt d) of antennas;
Step7: App displays the antenna parameters measurement results;
Step8: UAV returns and ends the measurement task.

III. Proposed System and Method

With the above prior knowledge and our exploration in measuring the antenna, an effective as well as reproducible system was designed for evaluating the antenna whose operation under a novel instance segmentation network was demonstrated in algorithm 1. To begin with, antenna video data is collected by UAV conducting random flying around the base station. Afterwards, the data obtained including data preprocessing, instance segmentation, and antenna parameter fitting and measurement, is to be processed by the designed novel instance segmentation network; at length, the APP layout in mobile terminal displays the content obtained by UAV and the measurement results of antenna parameters, and monitors the flight status of UAV in real time throughout the entire process. Experiments have proved the feasibility of this proposed system, while the market has verified its reproducibility and portability.

A. UAV Module

In this work, we define some certain flight parameters such as a flight radius of 5m and a flight angular velocity of 3°/s so as to capture antennas images of mobile communication base



(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 2. Samples of Antennascapes database: (a)-(d) from training set; (e)-(h) from testing set.

stations in various scenes and weathers for establishing a dataset named Antennascapes including 1,500 training images and 219 testing images is divided with some samples presented in Figure 2. VIA (VGG Image Annotator), a tool annotating images with open source, is adopted to annotate the front, side and front side antennas of training set which only have the annotation of antennas but the angle information. In other words, their down-tilt angle and other parameters cannot be acknowledged. Then professional workers are required to adjust down-tilt angle to a certain value (detailed in Table II) to construct a testing set which are also captured by UAV. There are 219 images in testing set which have no annotations but angle information. The creation of a testing set can not only verify the robustness and accuracy of the proposed instance segmentation algorithm but also demonstrate the effectiveness of the designed antenna parameter measurement system.

B. Instance Segmentation Module

With the continuous change of crowd density, the frequency of antenna parameter measurement and down-tilt adjustment is getting higher and higher. Regarding the challenges above, traditional methods are conducted with labor, while SOTA methods rely on expensive hardware equipment. This paper designed a novel instance segmentation network to solve these industry issues and the flowchart of our proposed algorithm can be seen in algorithm 2.

It is effective that the novel instance segmentation network is utilized to segment the antennas while a fitting and measuring module could get the antenna parameters using the pixel coordinate of segmented antennas. Figure 3 explains the architecture of proposed method. Among them, a basic network framework is built with the RetinaNet [27] based on FPN and resnet101. Masks for antennas in the proposed method are constructed through two simple parallel tasks, which is varied from traditional ones. Due to the utilization of fully connected layers in FCN to produce a group of prototype masks, which are identical in size to the primitive image, this turns into one parallel tasks.

The other one requires an extra prediction head added to detection branch so as to predict the mask coefficients' vectors.

Unfortunately, because Antennascapes precisely lack data with high possibility leading to training difficulties and

---

Algorithm 2:
Flowchart of designed instance segmentation network

Input:
A video of MCBS.
Output:
The number of antennas and their corresponding parameters.
Step1: Input a video of mobile communication base station;
Step2: Preprocessing module performs framing operations on the video and obtain antenna images;
Step3: The framed images are successively processed to generate prototype masks p and their corresponding mask coefficients m;
Step4: Generate new prototype masks p' using the most uncertain points extracted from the prototype masks and subsequently re-predicting the segmentation results;
Step5: Calculate the mask M for each antenna by multiplying m and p', and applying a threshold filter;
Step6: The antenna parameters are obtained by processing the mask M using a fitting and measurement module.
Step7: The value of n, representing the number of antennas, is established by applying a threshold to pixel coordinates, whereby the value of multiple-verifications for the empirical parameter is set to 50.
Step8: Number n of antennas, parameters of each antenna.

---

overfitting, transfer learning [28] suitable for small samples, which is proved by many studies is introduced to dispose of the problem appropriately.

Drawing inspiration from [17-18, 29], our design includes a dual attention module that refines the intermediate feature map at each convolutional block, utilizing both channel and spatial attention in turn. Since the assembly process is extremely lightweight, we can incorporate the dual attention module into every two convolutional layers seamlessly.

The severe imbalance of the method in positive and negative samples was solved by the Focal Loss, which successfully lessens the weight of a huge amount of simple negative ones. We implemented several empirical studies with the above prior knowledge to prove the effectiveness of our designed network. Simultaneously, comparative experiments were conducted be-

This article has been accepted for publication in IEEE Transactions on Artificial Intelligence. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAI.2023.3297991

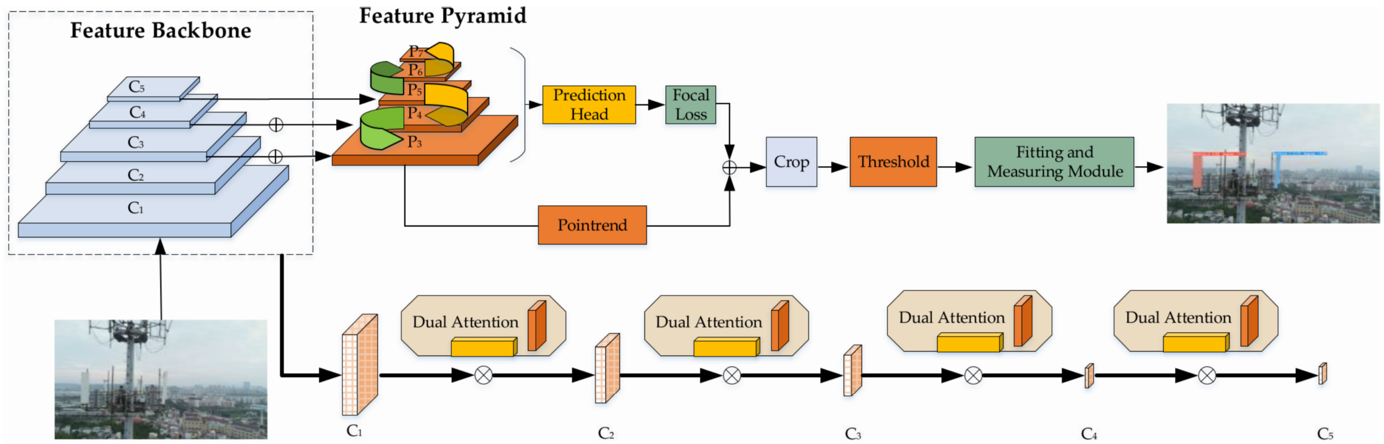


Fig. 3. The structure of proposed dual attention measuring network.

tween the proposed methods and existing ones on the same dataset, whose outcomes attest to the feasibility of lacking hardware in the designed instance segmentation network.

a) Instance Segmentation Module

Serial logic can be implemented using two-stage models that utilize re-pooling operations like ROI (Region of Interest) [10, 12]. However, alignment is required in [10] to map features to the bounding box, which makes it difficult to accelerate two-stage models. Despite efforts made by FCIS to parallelize re-pooling operations, it remains slow due to too many post-processing steps. In contrast, a full convolution model for segmenting instances in real-time [8] can achieve both precision and speed effectively.

As the first instance segmentation model that operates in real-time as a one-stage model. It directly integrates the mask branch into the one-stage detection algorithm, eliminating the need for ROI pooling. Learning from [8,10], the proposed instance segmentation network detailed in Figure 3 adopts one-stage object detection structures and a mask branch for the segmentation of antennas, which can speed up segmentation with details listed in [30-31]. The construction of masks for antennas is implemented with two simple parallel tasks, which is different from traditional instance segmentation tasks, and adds a dual attention mechanism module between the convolutional layers. In one of the parallel tasks, semantic vectors produced by fully connected layers in FCN are used to create prototype masks that match the size of the original image. In another parallel task, an additional prediction head is integrated into the one-stage object detection branch. This makes predictions that aid in the one-stage object detection task. This serves to figure out the mask coefficient's vectors. It can encode every anchor with the representation of instance in prototype space, which indicates the full utilization of the border of convolutional layers in engendering spatial coherent masks. Lastly, matrix multiplication is conducted to assemble with both spatial correlation preservation and one-stage model structure maintenance, featured with exceedingly efficient operation.

b) Transfer Learning

Transfer learning can reduce data from new domains by transferring features in low and middle levels from a related problem to a new one. Hence, with the adoption of transfer learning in a model which conducts image recognition with a huge number of images, it can alleviate the obsession due to not enough samples in an effective manner and realize the likely favorable performance. In this paper, transfer learning is capable of offsetting the dearth of antenna samples in the proposed instance segmentation network and measuring system.

The results obtained from the experiments effectively verify the function of transfer learning, whose theory is explained below. By utilizing knowledge taken from the source domain D_s and the domain corresponding to T_s and T_T, transfer learning enhances the ability of the object prediction function $f(T(.))$ in D_T, where it is assumed that either the domains or learning tasks are different ($D_s \neq D_T$ or $T_s \neq T_T$). For D_T, a limited number of antenna samples are used, while D_s consists of numerous images from various categories in MS COCO [32] or ImageNet [33]. Transfer learning is employed as a means of improving model performance as a result of the similarities between acquired models.

c) Dual Attention Mechanism

The instance segmentation network proposed in this study incorporates a dual attention mechanism module (DAMM) that is simple yet effective. The module makes use of an intermediate feature map to determine the attention weight along the channel and spatial dimensions, which is then applied to the original feature map for feature adaptation. Since the assembly process stands extremely lightweight, any CNN architecture can seamlessly integrate the DAMM with minor extra overhead. DAMM is a powerful attention mechanism for deep neural networks that enhances the representational power of convolutional blocks and its key advantage lies in its ability to adaptively recalibrate feature maps in a channel-wise and spatial-wise manner, thereby improving the model's capability on various computer vision tasks. Any basic CNNs is suitable for the end-to-end training of the dual attention mechanism. Both of these details can be referred to Figure 4.

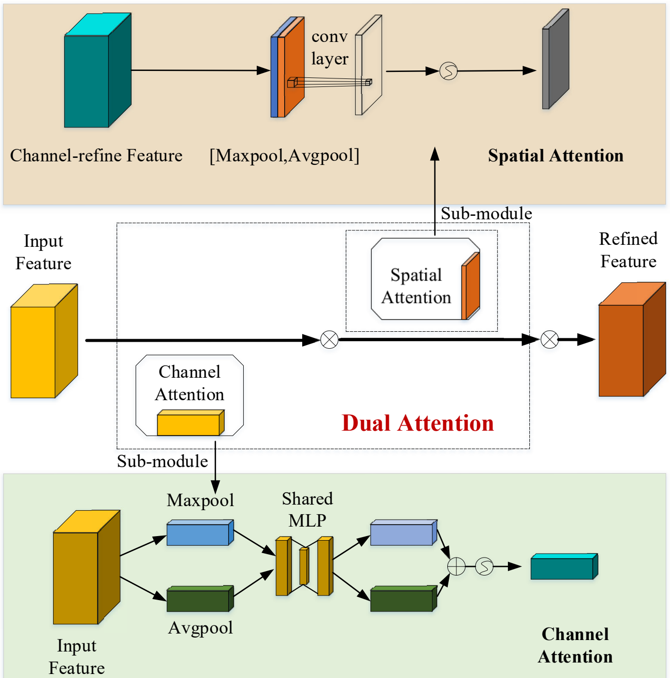


Fig. 4. illustrates the architecture of the dual attention module and its two sub-modules, namely the channel attention and spatial attention modules.

The dual attention module consists of two sub-modules: the channel attention module and the spatial attention module. These sub-modules utilize max-pooling and average-pooling outputs and share a common network. The spatial attention module combines similar outputs along the channel axis and maps them to a convolutional layer. The channel attention module highlights important features, where each channel represents a detector. By employing max pooling and global average pooling, the module gathers spatial features and captures diverse feature information, as following:

$$\begin{aligned} M_c(F) &= Sigmoid(f_{MLP}(Pooling_{Avg}(F)) + f_{MLP}(Pooling_{Max}(F))) \\ &= Sigmoid((W_1(W_0(F^C_{Avg})) + (W_1(W_0(F^C_{Max}))) \end{aligned} \tag{1}$$

The input to the model consists of a feature map F with dimensions $H \times W \times C$. Initially, $1 \times 1 \times C$ channel descriptors are derived from the feature map, employing both maximum pooling and global average pooling operations. Subsequently, these descriptors are passed through a standard two-layer neural network, wherein C/r and C denote the number of neurons in the first and second layers, respectively, while ReLU represents the activation function. Subsequently, attention coefficients M_c are generated by applying a sigmoid activation function to the obtained features. Multiplying these coefficients with the original feature F produces the new feature after scaling. Finally, the spatial attention module processes meaningful features using the equation shown below.

$$\begin{aligned} M_s(F) &= Sigmoid(f^{7\times7}([Pooling_{Avg}(F), Pooling_{Max}(F)])) \\ &= Sigmoid(f^{7\times7}([F^S_{Avg}; F^S_{Max}])) \end{aligned} \tag{2}$$

Both attention modules utilize feature F, which represents the spatial dimensions $H \times W \times C$ and the channel dimension C, as the input. Spatial attention module initiates with the introduction of an average and a max pooling of a channel dimension for two $H \times W \times 1$ channel description as well as for the splice of these two descriptions together according to channels. Subsequently, the weight coefficient is generated by passing through a 7×7 convolutional layer followed by a sigmoid activation function. Finally, multiplication is conducted on the feature F and the weight coefficient so as to obtain the new feature after scaling. Both of the attention modules are suitable for any basic CNNs regardless of how they are combined parallelly or sequentially. Experiments tell that sequential combination with channel attention mechanism placed ahead of spatial attention shows more favorable results. The two attention modules feature with the advantages of saving parameters and plug-and-play module into the current network architecture.

d) Focal Loss

As we know, in object detection and instance segmentation, a single image can generate thousands of candidate positions, but only a few of them actually contain objects, resulting in imbalanced categories. This category imbalance can significantly impact the optimization direction of the model and lead to unsatisfactory outcomes. To address this, we incorporated focal loss, a variant of conventional cross-entropy loss, into our training process. Focal loss serves to downweight easily classified samples, placing emphasis on the more challenging ones and prioritizing their learning. Focal loss achieves this by introducing a tunable hyperparameter called the focusing parameter, which weakens the contributions of easy examples and hence emphasizes more on the hard ones. In this way, the focal loss helps to alleviate the effect of category imbalance and leads to more accurate object detection and instance segmentation. Focal Loss is shown here:

$$FocalLoss(p_t) = -\alpha_t(1 - y_t)^{\gamma} \log(y_t) \tag{3}$$

When a sample is classified incorrectly, resulting in a small p_t, the modulation factor $(1 - y_t)$ will be prone to 1, standing for no change with the comparison to the original loss; then the weights of the well-classified samples are adjusted down. On the contrary, when y_t approaches to 1, it shows the correctness and easiness of the classification for samples, and meanwhile, when the modulation factor $(1 - y_t)$ is prone to 0, this sign represents the small contributions to the total loss. If $\gamma = 0$, we take the focal loss as the traditional cross-entropy loss. And γ shows positive correlation with the modulation coefficient. The focus parameter γ smoothly reduces the proportion of the weights of easily separated samples. Increasing γ enhances the influence of the modulation factor. It is proved by abundant experiments that 2 is the best value of γ.

e) Fitting and Measuring Module

Least squares method, is adopted to facilitate the calculation of the antenna parameters through studying the connection between the masked pixels. Least squares method weighs heavily in the analysis with a large amount of data. The 3/5-pixel value coordinates are obtained by removing the two

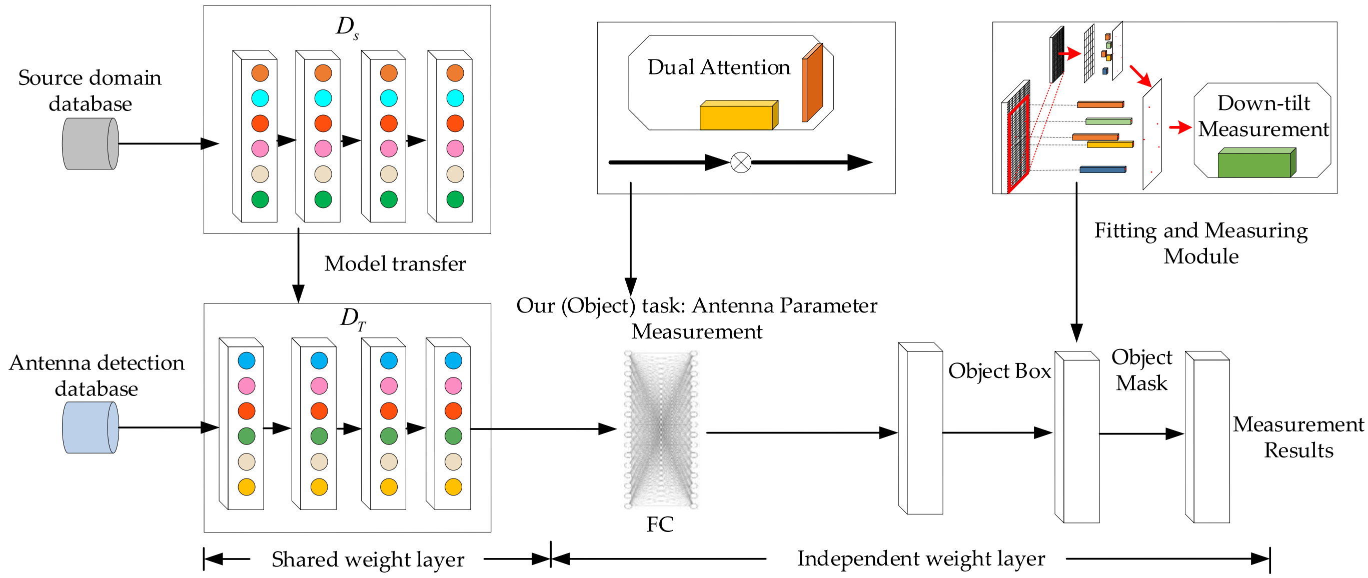


Fig. 5. The flowchart of the algorithm proposal.

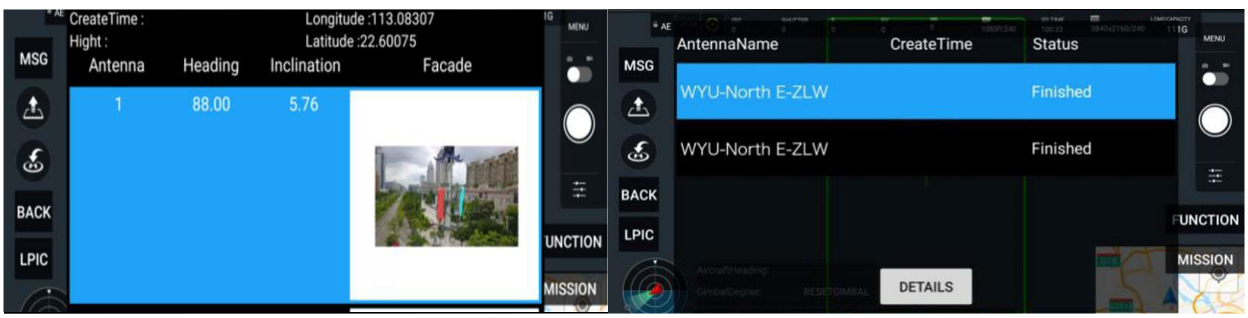


Fig. 6. The APP operational interface of the measurement system for antenna parameters .

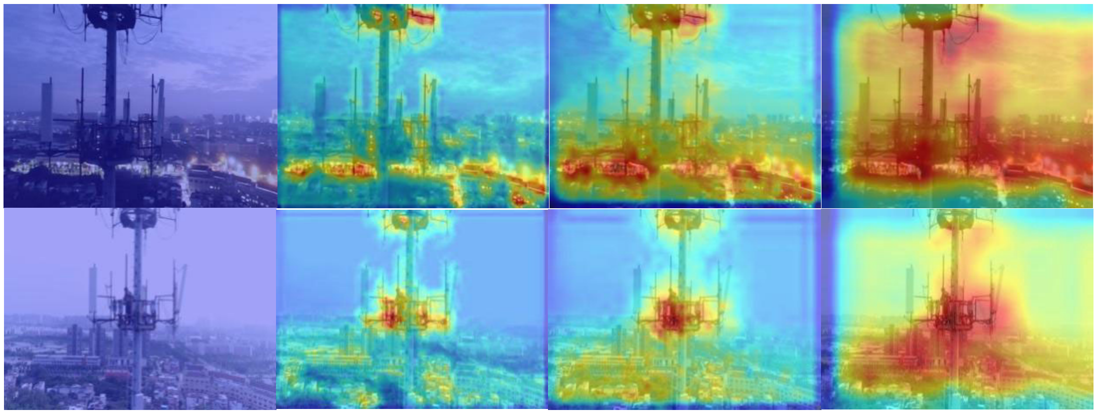
edge pixels on either side of the mask. These coordinates are specifically selected to accommodate the downward tilt angle of the antennas. Least square method is conducted for the purpose of fitting the selected pixels on two sides of antenna linearly to obtain a pair of slopes, and then calculate the arctangent of the pair of slopes to obtain the down-tilt angle of each side, finally the antenna down-tilt angle can be regarded as the average value of the antenna down-tilt angle of both sides. Based on the above prior knowledge and theoretical research, an overall description of the proposed instance segmentation algorithm is given. There are two stages in total. During training phase, transfer learning is for transferring the model trained by coco. After initializing the weights, the network is trained and the model is optimized using the training images from the constructed dataset. The dual attention module embedded between every two convolutional layers is used to make the network acquire more features from the target area, which should strengthen the model robustness. In the testing stage, the UAV flies around the base station to be tested to obtain antenna testing video, and transmits the obtained video to the server in real time through 5G transmission. And then the video is preprocessed by operations such as framing to obtain antennas testing images, and the optimal model is utilized to process these testing images to obtain the mask of each antenna. Finally, the fitting and measurement module is utilized to analyze each mask and extract the parameters of individual antennas. The overall framework of the proposed system is depicted in Figure 1, showcasing the key components and their interactions. Additionally, Figure 5 provides a comprehensive flow chart illustrating the algorithm designed for the system. These figures collectively present a visual representation of the system's architecture and algorithmic workflow.

C. APP Module

With the popularization of mobile Internet and technological development and progress, the focus of product has become one of the key issues that users pay attention to. To this end, we have developed an Android Application (APP) to perfect the system for measuring the antenna parameters in mobile communication base station. The main functions of this APP are to control and monitor the flight status of UAV, display the flight parameters of UAV and the measurement results of each antenna parame- ters to ensure a better user experience. With the app being put into use, facts have proved that it is effective for improving work efficiency. The operation interface of part of the APP is shown in Figure 6. The figure shows that our proposed system is performing antenna parameter measure- ment, and the measurement result has also been returned to the APP interface.

IV. Experimental Analysis

We arrange a Core i9 CPU computer with 64G memory and 2 Titan RTX graphic cards with 48G in the experiments. Further-



(a) (b) (c) (d)

Fig. 7. Outputs from every dual attention module on resnet50, in which the former two images are at down-tilt angles of 5 and 12 while (a)-(d) separately refer to the attention maps between every two convolutional layers.

TABLE I
EXPLANATION OF ANTENNASCAPES

	Number	Annotation	Angel Information
Training set	1500	√	
Testing set	219		√

TABLE II
DISTRIBUTION OF THE TESTING SET

Testing set including information on angel investors									
Angel	1°	4°	5°	6°	7°	8°	9°	12°	15°
Number	13	25	22	30	25	23	22	39	20

TABLE III
RESULTS OF COMPARING VARIOUS TRANSFER MODELS

	MS COCO [32]		ImageNet [33]	
	T	F	T	F
Segmentation accuracy	**87.73%**	72.90%	86.91%	72.90%
Fitting accuracy	**70.83%**	59.68%	70.09%	59.68%

TABLE IV
A COMPARISON TABLE IS PRESENTED FOR THE RESULTS OBTAINED FROM VARIOUS FEATURE EXTRACTION NETWORKS, WITH AND WITHOUT THE APPLICATION OF DUAL ATTENTION MECHANISM.

	Resnet101(%)		Resnet50(%)		Darknet53(%)	
	T	F	T	F	T	F
Segmentation accuracy	63.54	63.22	**72.91**	61.73	68.71	67.07
Fitting accuracy	51.63	38.83	**59.68**	37.83	47.59	35.36

more, we also adopt pytorch1.1.0 and Ubuntu18.04 operating system. A dataset called Antennascapes which is full of natural scenes from low-altitude areas is established consis- ting of 1500 training images and 219 testing images. Some images from the database are shown in Figure 3. Because all the images are shot by UAV flying around the station, they all possess higher resolution than MS COCO [32] and ImageNet [33] and more favorable pixel preciseness for GT instance segmentation. The training set consists of 1500 artificially and accurately labeled images, while the testing set consists of 219 images that were manually labeled at specific angles (1-degree, 4-degree, 5-degree, 6-degree, 7-degree, 8-degree, 9-degree, 12-degree, and 15-degree) to test the precision of the experiment. Table I and Table II show the distribution of the Antennascapes.

A. *Analysis of Transfer Learning*

For the positive correlation between the training samples quantity and the trained model accuracy, we introduce transfer learning method to make up for inadequate samples. The weights from models trained on MS COCO and ImageNet are used for initialization to prevent random weight initialization. The model only uses the resnet50 and FPN modules to transfer weights from these models. Table Ⅲ displays the segmentation and fitting accuracies achieved by this approach. The "T" column stands for the designed network with transfer learning while "F" column refers to the designed network without transfer learning. After the comparison, it is known that a more favorable accuracy in both segmentation and fitting is realized in MS COCO dataset model through the weight initialization of MS COCO dataset model, compared with that in ImageNet dataset model. Apparently, with the assistance of transfer learning, a favorable growth can be observed in both of the accuracies. These obvious performances can be convincing that transfer learning can tremendously raise the accuracy in segmentation among tasks with small samples.

B. *Analysis of Dual Attention Mechanism*

We train all models at the batch size of 16 on two GPUs transferring MS COCO pre-trained model which has been experimentally proved that the performance is better than when using ImageNet pretrained model. The pretrained batch specification is not frozen and an additional BN layer is not added since the arranged batch size is able to adequately meet the standard of batch norm. The model is trained using SGD with an initial learning rate of 10-4 for 10000 iterations. The learning rate is divided by 10 at iterations 2000, 7000, 8000, and 8500, with a momentum set at 0.9 and a score threshold of 0.247. Weight decay of $5e^{-5}$ is employed, and all data augmentations are carried out in SSD. The processing of our 1500 training images costs 10 to 12 days.

Figure 3 shows the distribution of section 3, in which every two convolutional layers are inserted with the dual attention (channel attention ahead of spatial attention). We evaluate the effectiveness of our dual attention module by combining various feature extraction networks (resnet101, resnet50, darknet53) and the FPN network. The evaluation is conducted under the pre-training model of MS COCO, and the specific comparison results are presented in Table IV. In this comparison, the segmentation accuracy is measured as the ratio

This article has been accepted for publication in IEEE Transactions on Artificial Intelligence. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAI.2023.3297991

Fig. 8. The testing image contains data for an antenna with an actual down-tilt angle of 12 degrees. The measured down-tilt angle is recorded as 12.03 degrees.

TABLE V
PERFORMANCE COMPARISON OF BACKBONE MODELS USING MULTIBOX LOSS

Backbone	Segmentation accuracy	Fitting accuracy	Time(s)
Resnet101	63.54%	51.63%	0.15
Darknet53	68.71%	47.59%	0.13
Resnet50	**72.91%**	**59.68%**	**0.14**

TABLE VI
COMPARISON OF RELATED METHODS ON THE SAME DATASET

Loss	Segmentation accuracy	Fitting accuracy	Time(s)
Multibox[31]	72.91%	59.68%	0.14
Triplet[34]	75.24%	67.48%	0.12
Center[35]	100%	80.16%	0.14
Circle[36]	100%	80.25%	0.12
Focal	**100%**	**85.17%**	**0.13**

TABLE VII
MEAN ERROR AND VARIANCE ANALYSIS

GT Angle (degree)	1	4	5	6	7	8	9	12	15
Mean error	0.42	0.34	0.75	0.26	0.09	0.19	0.09	0.17	0.35
Variance	0.09	0.46	0.17	0.15	0.43	0.30	0.06	0.14	0.12

TABLE VIII
COMPARISON OF THE PROPOSED METHOD WITH OTHER EXISTING METHODS ON THE SAME DATASET

Method	Backbone	Segmentation/ Detection accuracy	Fitting accuracy	Time (s)
Faster R-CNN [12]	Resnet101	89.85%	-	3.08
YOLOv3[30]	Darknet53	92.72%	-	0.01
YOLOX-S [37]	Darknet53	85.95%	-	0.01
YOLOX-M [37]	Modified CSP	94.48%	-	0.22
DETR [38]	Resnet50	91.22%	-	0.59
CenterNet [39]	DLA-34	98.80%	-	0.03
R-CenterNet	DLA-34	98.80%	74.33%	0.04
S^2ANet [41]	Resnet50	90.01%	61.33%	0.06
Mask R-CNN [10]	Resnet101	99.44%	58.20%	5.95
Ours	Resnet50	**100%**	**85.17%**	**0.13**

of segmented non-back antennas to all non-back antennas. What is more, most often, less than 1-degree deviation between measured and real value can be overlooked in practice. Therefore, we can define fitting accuracy as the proportion of the amount of segmented front-side antennas, whose down-tilt angles are evaluated within 1° compared to the real down-tilt angles of the quantity of detected front-side antennas.

The attributes "T" and "F" differ depending on the presence of a dual attention mechanism module within the network. In Table III, it is written with the best segmentation accuracy at 72.91% and fitting accuracy climbs to 59.68%. This performance belongs to resnet50 with dual attention. We compare the figures also from resnet50 but without attention mechanism, which are listed in the next column. The figures present a great improvement in the accuracy of segmentation and fitting. Attention maps engendered from dual attention module in resnet50 with FPN can be referred to Figure 7. The theory in these images is that the red color positively correlates to the contribution made to the final prediction result while the blue color does the opposite.

C. *Analysis of Focal Loss*

We compare segmentation accuracy, detection accuracy, fitting accuracy and time spending of proposed method with other current Loss Functions including Multibox Loss [31], Triplet Loss [34], Center Loss [35], and Circle Loss [36]. The training strategy and all data augmentations in this work are based on SSD, so is Loss function. Therefore, under the premise of controlling the same training strategy and evaluation criteria, Multibox Loss is introduced to the proposed method to conduct experiments on different backbones which are detailed in Table V to find the best performing backbone on our instance segmentation task. It is proved that when the backbone is Resnet50, the test evaluation criteria are the best. Then, the feature extraction network was unified as resnet50, and we conducted comparative experiments with different loss functions. Table VI describes that when focal loss is adopted to our network, the segmentation accuracy and fitting accuracy perform best. Finally, Comparative experiments between methods which are detailed in Table VI are conducted with the same dataset.

D. *Analysis of Fitting and Measuring Module*

In Figure 3, the Fitting and Measuring module is imposed to fit and measure parameters of mask of each antenna output by the instance segmentation network. This work requires professional workers to adjust the down-tilt angle for daily use and compares it with the results of our proposed algorithm within the allowable error range so that the Fitting and Measuring module can verify itself. In measuring the antenna parameters, measurement errors within one degree are allowed.

An image in testing set with actual down-tilt angle as 12 degree is processed by our proposed network to get the antennas masks, and the Fitting and Measuring module is adopted to measure the down-tilt angle. Figure 8 clearly shows that the measurement result is 12.03 degree. It is surprised to find that the measurement result of our proposed system is not only infinitely close to the actual angle but also exceeds most of the manual random measurement results.

Our method demonstrates inspiringly satisfying performance in the discussed field compared to current advanced methods, which indirectly proves and emphasizes the potency and essentialness of dual attention mechanism by transfer learning. Table VIII provides enough data to choose the best-performance model through the comparison of the deliberated and the actual angles. Since the error value of each group is less than 1 degree, it conforms to the requirement of the industry. Meanwhile, we have conducted a mean error and variance analysis of the proposed method. Among them, error is the numerical value of the deviation between the true and the measured value. From Table VII, it can be reached that the proposed system has well performance of good effectiveness and stability and the measurement results are not volatile from mean error and variance analysis. It is surprising that fluctuation can be barely observed at the same angle according to the error analysis in testing set. This attests to the outstanding stability and reliability of the proposed system in this work.

E. Comparisons of other SOTA Methods

In Table VIII, the proposed approach is compared with the current most advanced models, and Resnet50 is used as a feature extraction network. The evaluation metrics applied contain detection accuracy, segmentation accuracy, fitting accuracy and time cost. It has been shown that our approach surpasses all other methods in these aspects. Notably, since the object detector head outputs an axis-aligned bounding box, only the detection accuracy can be compared, and the fitting accuracy is not reported. To enhance the orientation prediction, we incorporated a branch for predicting the rotation angle of the bounding box in CenterNet's [39] head, which transformed it into a rotating object detector. The accuracy rate of the inclination angle measurement for the test set was 74.33%, whereas S^2ANet [40] achieved a measurement accuracy of 61.33% following the same bounding box rotation angle prediction principle. Nonetheless, experiments further demonstrate that the rotating object detector can only offer a rough estimation of the object's orientation and cannot be applied to the precise angle measurement scene since the inconsistency of angular boundaries is a long-standing bottleneck and challenge of the orientation detector. Conversely, the proposed method attained a detection accuracy of 100% and a fitting accuracy of 85.17% in the test set, and the average measurement time for all segmented antennas in the image was 0.13s, surpassing the accuracy of segmentation and fitting. Verification by some existing deep learning-based detection methods and instance segmentation algorithms has shown that our method is effective.

V. Conclusion

We present an innovated measurement method for antenna parameters measurements in mobile communication base station, which replaces traditional measurement methods with human labor involved. A novel instance segmentation framework is designed with the employment of dual attention module. The combination of fitting module and Focal loss can accelerate the evaluation of antenna parameters. After complete experiments, we can acknowledge that the adoption of proper training strategies can produce satisfying performance in the automatic measurement of antenna parameters. Comparisons conducted among current instance segmentation algorithms and detection algorithms derived from deep learning methods and our method verify the predominant performance of our method in rapid recognition and favorable fitting accuracy. Moreover, our proposed method stands with less manual interference, better efficiency and higher safety than the traditional methods. At length, it is forecast that our proposed method for antenna parameters measurement will occupy a significant place in 5G era featured with high bandwidth and concurrency and low latency.

However, there are still many research challenges in this work. When deploying the model in real industrial scenarios, it relies on high-configured servers, and the remote server is affected by the network signal from the local end. To achieve long-term fully autonomous AI measurements, the algorithm needs to be lightweight and directly integrated into the drone flight system to break free from the constraints of servers and devices. Furthermore, the proposed method can currently only measure the down-tilt angle of antennas. The automatic measurement of more antenna parameters, like azimuth angle and weak coverage areas, has yet to be resolved.

References

[1] LeCun. Y, Bengio. Y, and Hinton. G, "Deep Learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May, 2015.

[2] Krizhevsky, Alex, et al., "ImageNet Classification with Deep Convolutional Neural Networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, June, 2017.

[3] Lee, Woongsup, et al. "Deep Learning Based Transmit Power Control in Underlaid Device-to-Device Communication," *IEEE Systems Journal,* vol. 13, no. 3, pp. 2551–2554, Sep, 2019.

[4] Kun. X, Guo. T, Zhou. Y, et al., "Antenna Mechanical Pose Measurement Based on Structure from Motion," *ZTE Communications*, vol. 4, no. 16, pp.38-45, 2020.

[5] Pedras. V, Sousa. M, Vieira. P, et al., "Antenna tilt optimization using a novel QoE model based on 3G radio measurements", *in Proceedings of 20th International Symposium on Wireless Personal Multimedia Communications (WPMC),* Indonesia, Yogyakarta, 2017, pp. 124-130.

[6] Hafiz. A, and Ghulam. M, "A Survey on Instance Segmentation: State of the Art," International Journal of Multimedia Information Retrieval, vol. 3, no. 3, pp. 171–189, June. 2020.

[7] Huang, Zhaojin, et al., "Mask Scoring R-CNN." *in Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),* California, USA, 2019, pp. 6409–6418.

[8] Bolya. Daniel, et al., "YOLACT: Real-Time Instance Segmentation," *In Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV),* Korea, Seoul, 2019, pp. 9157–9166.

[9] Bai M, Urtasun R Deep Watershed Transform for Instance Segmentation. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 21-26 July 2017 2017. pp 2858-2866.

[10] He. K, Georgia. G, Piotr. D, "Mask R-CNN," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 42, no. 2, pp. 386-397, Feb. 2020.

[11] Girshick, Ross. "Fast R-CNN," *in Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV),* Santiago, Chile, 2015, pp. 1440–1448.

[12] Ren. S, He. K, Girshick. R, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 39, no. 6, pp. 1137-1149, June. 2017.

[13] Shelhamer. E, J, and Darrell. T, "Fully Convolutional Networks for Semantic Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 39, no. 4, pp. 640-651, Apr. 2015.

[14] T. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan and S. Belongie, "Feature Pyramid Networks for Object Detection," *in proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* Honolulu, HI, 2017, pp. 936-944.

[15] Vaswani. A, Shazeer. N, Parmar. N, *et al.,* "Attention Is All You Need," *in Proceedings of the International Conference on Neural Information Processing Systems,* Vancouver, Canada, 2017, pp. 5998–6008.

[16] Wang X, Girshick R, Gupta A, et al., "Non-local neural networks," *in proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),* Salt Lake City, UT, 2018, pp.7794-7803.

[17] Woo. S, Park. J, Lee. J, Kweon. I, "CBAM: Convolutional Block Attention Module," *in Proceedings of the European Conference on Computer Vision (ECCV),* Munich, Germany, 2018, pp. 3–19.

[18] J. Fu, et al., "Dual Attention Network for Scene Segmentation," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition,* Long Beach, CA, USA, June 2019, pp. 3141-3149.

[19] T. Lin, P. Goyal, R. Girshick, K. He and P. Dollár, "Focal Loss for Dense Object Detection," in *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 42, no. 2, pp. 318-327, Feb. 2020.

[20] Romdhane T F, Pr M A. "Electrocardiogram heartbeat classification based on a deep convolutional neural network and focal loss," *Computers in Biology and Medicine,* vol. 123, pp. 103866-103887, Aug, 2020.

[21] Yun P, Tai L, Wang Y, et al., "Focal loss in 3d object detection," *IEEE Robotics and Automation Letters,* vol. 4, no. 2, pp. 1263-1270, Apr, 2019.

[22] Li D, Wen G, Kuai Y, et al., "Robust visual tracking with channel attention and focal loss," *Neurocomputing,* 401: 295-307, Aug, 2020.

[23] Samaras S, Diamantidou E, Ataloglou D, et al., "Deep learning on multi sensor data for counter UAV applications—A systematic review," *Sensors,* vol. 19, no. 22, pp. 4837-4871, Nov, 2019.

[24] Barbedo J G A, Koenigkan L V, Santos T T, et al., "A study on the detection of cattle in UAV images using deep learning," *Sensors,* vol. 19, no. 24, pp. 5436-. Dec, 2019.

[25] Kyrkou C, Plastiras G, Theocharides T, et al., "DroNet: Efficient convolutional neural network detector for real-time UAV applications," *In Proceedings of 2018 Design, Automation & Test in Europe Conference & Exhibition (DATE),* Dresden, Germany, 2018, pp. 967-972.

[26] Walker O, Vanegas F, Gonzalez F, et al., "A deep reinforcement learning framework for uav navigation in indoor environments," *In Proceedings of 2019 IEEE Aerospace Conference.* Yellowstone, USA, 2019, pp. 1-14.

[27] Zhang, Hongkai, et al., "Cascade RetinaNet: Maintaining Consistency for Single-Stage Object Detection," *in 2019 British Machine Vision Conference (BMVC),* Cardiff University, United Kingdom, Sep, 2019, pp. 227-238.

[28] Pan. S, and Yang. Q. "Transfer Learning," *IEEE Transactions on Knowledge and Data Engineering,* vol. 22, no. 10, pp.1345 – 1359, Oct. 2010.

[29] Mukherjee, Anandarup, et al. "MiND: Mind Networked Device Architecture for Attention-Gated Ambient Assisted Living Systems." *IEEE Systems Journal,* vol. 14, no. 1, pp. 1325–1332, Mar, 2020.

[30] Redmon. J, Ali. F, "YOLOv3: An Incremental Improvement," *ArXiv Preprint ArXiv*:1804.02767, Apr. 2018.

[31] Liu. W, Anguelov. D, Erhan. D, et al., "SSD: Single Shot Multi-Box Detector," *in Proceedings of European Conference on Computer Vision (ECCV),* Amsterdam, Netherlands, 2016, pp. 21–37.

[32] Lin. T, Maire. M, Belongie. S, et al., "Microsoft COCO: Common Objects in Context," *in Proceedings of European Conference on Computer Vision (ECCV),* Zurich, Switzerland, 2014, pp. 740–755.

[33] Jia. D, Wei. D, Socher. R, et al., "ImageNet: A Large-Scale Hierarchical Image Database," in Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Florida, USA, 2009, pp. 248–255.

[34] F. Schroff, D. Kalenichenko and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," *in Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* Boston, MA, 2015, pp. 815-823.

[35] Wen, Yandong, et al. "A Discriminative Feature Learning Approach for Deep Face Recognition," *in Proceedings of European Conference on Computer Vision (ECCV),* Amsterdam, Netherlands, 2016, pp. 499–515.

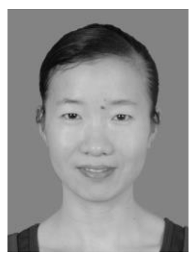
[36] Sun, Yifan, et al. "Circle Loss: A Unified Perspective of Pair Similarity Optimization," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),* online, 2020, pp. 6398–6407.

[37] Ge, Zheng, et al., "Yolox: Exceeding Yolo Series in 2021," *in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021.

[38] Carion, Nicolas, et al., "End-to-End Object Detection with Transformers," *in Proc. Eur. Conf. Comput. Vis. (ECCV)*, Glasgow, UK, Aug. 2020, pp. 213–229.

[39] Zhou, Xingyi, et al., "Objects as Points," *in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Long Beach, CA, USA, Jun. 2019.

[40] Han, Jiaming, et al., "Align Deep Features for Oriented Object Detection," *IEEE Trans. Geosci. Remote Sens*., 2021.



Ying Xu received the B.S. and M.S. degrees in Automation, and Control Engineering from Wuhan University of Science and Technology in 2004 and 2008, respectively. She also received her Ph.D. in South China University of Technology in 2013.

She joined Wuyi University in 2008, and her research interests include: intelligent signal processing and pattern recognition.

Qirui Ke is currently pursuing his Master in Department of Intelligence Manufacturing, Wuyi University. He received his B.S. degree at Hubei University of Arts and Science in 2018. His research interests include: pattern recognition and automatic measuring.

Ziyi Jiang is currently pursuing his Master in Department of Intelligence Manufacturing, Wuyi University. He received his B.S. degree in Wuyi University, Jiangmen in 2019. His research interests include: image processing, pattern recognition and automatic measuring.

Yikui Zhai is an Associate Professor at Wuyi University, Guangdong, China. He was a visiting scholar at the University of Milan, Department of Computer Science, since 2016. He received his Ph. D degree in Signal and Information Processing at Bei hang University in June 2013. He received the Bachelor and Master degree in Optical Electronics Information and Communication Engineering, and Signal and Information Processing from Shantou University, Guangdong, China in 2004 and 2007 respectively.

Since October 2007, he has been working at Department of Intelligence Manufacturing, Wuyi University, Guangdong, China, and his research interests include: image processing, deep learning and pattern recognition.

Angelo Genovese received the Ph.D. degree in computer science from the Università degli Studi di Milano, Crema, Italy, in 2014. He has been a postdoctoral Research Fellow in computer science with the Università degli Studi di Milano since 2014.

He has been a Visiting Researcher with the University of Toronto, Toronto, ON, Canada. Original results have been published in over 30 papers in international journals, proceedings of international conferences, books, and book chapters. His current research interests include signal and image processing, three-dimensional reconstruction, computational intelligence technologies for biometric systems, industrial and environmental monitoring systems, and design methodologies and algorithms for self-adapting systems.

Dr. Genovese is an Associate Editor of the Journal of Ambient Intelligence and Humanized Computing (Springer).

Vincenzo Piuri (Fellow, IEEE) received the M.S. and Ph.D. degrees in computer engineering from Politecnico di Milan, Milan, Italy, in 1984 and 1988, respectively.

He was the Department Chair with the University of Milan, Milan, from 2007 to 2012, where he has been a Full Professor since 2000. He was an Associate Professor with Politecnico di Milan from 1992 to 2000, a Visiting Professor with The University of Texas at Austin, Austin, TX, USA from 1996 to 1999, and a Visiting Researcher with George Mason University, Fairfax, VA, USA, from 2012 to 2016. He founded a startup company, Sensure srl, Bergamo, Italy, in the area of intelligent systems for industrial applications (leading it from 2007 to 2010) and was active in industrial research projects with several companies. His main research and industrial application interests are intelligent systems, computational intelligence, pattern analysis and recognition, machine learning, signal and image processing, biometrics, intelligent measurement systems, industrial applications, distributed processing systems, Internet-of-Things, cloud computing, fault tolerance, application-specific digital processing architectures, and arithmetic architectures.

Dr. Piuri is an ACM Fellow.

Fabio Scotti (Senior Member, IEEE) received the Ph.D. degree in computer engineering from the Politecnico di Milan, Milan, Italy, in 2003.

He has been an Associate Professor of computer science with the Universitá degli Studi di Milan, Crema, Italy, since 2015. His original results have been published in more than 100 articles in international journals, proceedings of international conferences, books, book chapters, and patents. His research interests include biometric systems, machine learning and computational intelligence, signal and image processing, theory and applications of neural networks, 3-D reconstruction, industrial applications, intelligent measurement systems, and high-level system design.

Dr. Scotti is an Associate Editor with the IEEE Transactions on Human-Machine Systems and Soft Computing (Springer). He has been an Associate Editor with the IEEE Transactions on Information Forensics and Security, and a Guest Coeditor for the IEEE Transactions on Instrumentation and Measurement.